DOCUMENT RESUME

ED 097 017 IR 001 199

TITLE	Sharper Tools for Better Learning.
INSTITUTION	National Association of Secondary School Principals, Washington, D.C.
PUB DATE	73
NOTE	43p.
AVAILABLE FROM	The National Association of Secondary School Principals, 1904 Association Drive, Reston, Virginia 22091 ($2.00)
EDRS PRICE	MF-$0.75 HC Not Available from EDRS. PLUS POSTAGE
DESCRIPTORS	*Curriculum Evaluation; Equipment Evaluation; *Evaluation Criteria; Evaluation Methods; *Instructional Improvement; Instructional Innovation; *Instructional Materials; Instructional Materials Centers; Instructional Media; Instructional Technology; *Media Selection; State of the Art Reviews; Textbook Evaluation

ABSTRACT

Procedures and practices that lead to better selection of instructional materials for schools are described in this report, which also highlights the current state of curriculum evaluation. The reationale for evaluating instructional materials is explored for its real meaning and purpose. Selection committees, the usual means for instructional material selection in schools, are explained, and the advantages to be obtained by different format, organization, membership, and training are shown. The critical phases in selection, namely, establishing selection criteria, examination and review of materials, and pilot use and testing of materials, are described. Finally, the expectations that can be made of educational research toward selection of instruction materials are analyzed. (WH)

ED 097017

SHARPER TOOLS FOR BETTER LEARNING

BEST COPY AVAILABLE

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

NASSP

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

NATIONAL ASSOCIATION OF SECONDARY SCHOOL PRINCIPALS
1904 Association Dr., Reston, Va. 22091

Copyright 1973

THE NATIONAL ASSOCIATION OF SECONDARY SCHOOL PRINCIPALS

1904 Association Dr., Reston, Va. 22091

ISBN: 0-88210-048-3

$2 per copy. 10-24 copies, 20% off; 25-99 copies, 30% off; 100 or more copies, 35% off. **All orders totaling $10 or less must be accompanied by full payment.**

Contents

Foreword v

Why Sharper Tools? 1

Better Materials, Not Just More 3

The Selection Committee 7

Establishing Criteria 9

Examination and Review 17

Pilot Use and Testing 25

What Can We Expect Others To Do? 29

Appendix A 35

Appendix B 39

Bibliography 41

Foreword

Instructional materials are an educator's tools. They are used in classrooms throughout the country to improve teaching and effect learning, and their cost represents a good portion of every school's budget.

At one time, instructional materials meant textbooks almost exclusively. Today the term includes an abundance of audio and visual materials, and there's no apparent end to what can be developed as educational technology progresses.

Selecting instructional materials, always an important task, has become increasingly difficult as more and more materials have become available. *Sharper Tools for Better Learning* acknowledges this difficulty and provides valuable information to the administrator searching for ways to improve the selection process.

We are grateful to NASSP's Committee on Educational Technology for initiating this monograph. Committee Chairman Warren Koch, Consultant Wesley Walton, and members of the committee offered invaluable assistance in revisions of the manuscript. We also thank Kenneth Komoski, president of Educational Products Information Exchange, (EPIE), for his cooperation in the project and Betty Preston, who prepared the original manuscript.

Owen B. Kiernan
Executive Secretary, NASSP

Why Sharper Tools?

Schools today are demanding better and sharper educational tools—tools that will be more effective than traditional ones in meeting the needs of individualized student learning, greater teacher accountability, and additions to the curriculum.

We are only now beginning to realize that independent learning materials, oriented to the individual, differ in many ways from the teacher-dependent materials, oriented to the class, which still predominate in schools. We are gradually discovering that the inadequacies of traditional materials cannot be compensated for by adding new and different materials merely because they are new and different.

We are also very slowly coming to the realization that, in the interest not only of education but also of economy, materials must demonstrate the degree of their effectiveness before we buy them, not after a somewhat haphazard on-the-job trial in our schools.

When curricula were more or less standard, when available materials were less numerous and more alike than different, and when what we wanted was something to cover the necessary ground for everyone, with some hints at "outside projects" for the very gifted, there was less chance of selecting totally inappropriate materials. As a result, we got into the habit of considering examination and review by experienced teachers as the best possible way to select instructional materials. This method now is failing us.

Today some schools are beginning to use materials experimentally with pilot groups of students before making the materials a part of the curriculum; or to ask for evidence, based on measured student performance, that a piece of material will do what it is supposed to do.

Producers of instructional materials rarely provide

real evidence of product capability in response to requests for it. They do little field testing and almost no developmental testing (trying materials with students and revising them in light of their performance). Post-publication feedback is almost unheard of. Producers too often rely on expert opinion as schools have done, since we have demanded no more of them.

> "We have about 160 salesmen and consultants who report back what they pick up in the field; that's really our field testing."
>
> *Vice President, Editorial, of a major publishing company*

So we find ourselves applying inadequate selection procedures to inadequately developed educational materials.

The importance of materials selection becomes clear when we realize that around 75 percent of a student's classroom time and 90 percent of his homework time is spent in using instructional materials. Schools spend too little time selecting teaching materials and only about three percent of their operating budgets buying them.

The process for selecting better instructional tools has to be refined. School principals must take the responsibility for persuading users and producers that instructional tools must be better selected and better developed. This monograph gives the necessary information on what to do and how to do it.

Better Materials, Not Just More

"Everything that *can* be done *must* be done to get the education industry to fulfill its potential; not by providing the schools with endless—but in the end trivial—options, but by supplying effective alternatives for individual learners. The potential impact of its products is enormous. These products contain materials that introduce (or fail to introduce) skills, concepts, facts, and understanding into the minds of 50 million young Americans for twelve crucially important years. No industry in the country produces products of greater importance or potential."

The Director of EPIE, testimony before a House subcommittee, May 11, 1971

Some 200,000 items of instructional materials — books, films, tapes, kits, etc. — are on the market today. This is 20 times more than two decades ago!

Why this proliferation of materials? Let's look at the reasons.

- First and most dramatic, the introduction of several new media. The textbook is no longer the sole source of information, sometimes not even the primary one. In some cases it may have been supplanted unwisely.
- Similarly, the introduction of several new learning topics. A noted education historian once observed that other countries met social problems with revolutions, but that America simply added a course to the curriculum. Thus, we now have courses in black studies, drug abuse, and ecology.
- Then, the evolution of new approaches to teaching. "Discovery" and "using a problem-solving approach" are examples, as are changes like teaching machines and the Initial Teaching Alphabet.

- Also, the many changes in course content. Some changes were necessitated by advances in knowledge, but others came out of a realization that what we had been doing was not doing the job. (It was in these latter cases that the beginnings of a scientific approach to the development of all instructional materials, not just tests and some programed materials, came about.)
- Finally, the increase in purchasing power made possible by federal legislation. While some federal funds were used to improve materials, others served merely to increase the demand for materials.

The proliferation of materials has made selection of good materials difficult. Not only do we have many more materials to choose from, but we are also beginning to establish new types of learning environments geared to individual learning and to realize that the quality of materials put into the hands of learners is critical to the success of individualized programs. We see the need for pinpointing very specific goals, and for finding materials to meet those goals.

"Accountability is negative in that it draws attention to what is just enough. Schools have higher aims. They cannot be wholly accountable unless they are much more than merely accountable."

"The Accountability Notebook" prepared by Center for Instructional Research and Curriculum Evaluation (CIRCE), University of Illinois

The question of our own accountability, for students' learning and for public funds, sharpens our perception not only of the inadequacies of the materials we have used but also of the inadequacy of the information we have relied on to help us select those materials. We must change our own well-established, comfortable behavior.

What Can WE Do?

- We must get more people into the act: parents, administrators, and students — not just teachers.
- We must state explicitly our criteria for selection: cooperatively developed lists — not just assumptions that all seek the same ends.
- We must expand our basic information sources: comprehensive listings and collections — not just advertisements and producers' mailings.
- We must search out more kinds of information: professional analyses, user reports, empirical evidence of effectiveness — not just producers' or authors' statements.
- We must refine our techniques of examination and review: comprehensive check lists, rating scales, and questionnaires — not just unstructured comments.
- We must add scientific evaluation of materials: trial uses empirically evaluated and systematic feedback once materials are adopted and put to use — not just opinion and impression.

The Selection Committee

A common misapprehension is that the best way to obtain useful and appropriate educational materials is to ask an experienced teacher to select them.

Although based on misapprehension, this method certainly is not the worst. But the selection job should be shared. Opinions of students, parents, and administrators should be added to those of teachers — several teachers, not just one. Such a selection committee ensures that all concerned points of view will be heard and that all concerned groups will be informed about the availability and purposes of materials. Such a group also provides the manpower needed for the work of searching, sifting, and evaluating.

Conduct Training Sessions

Typically a selection committee sets up its own regulations and operating procedures, under the leadership of the person with final responsibility — usually the principal.

Members of selection committees generally have little or no special preparation in evaluation techniques. Time, therefore, should be spent in training the committee on how to establish and apply criteria, and how to devise valid check lists and rating scales. Time should also be spent in evaluating a few items for practice. Specific training may be needed in simple evaluation techniques. The evaluation specialist serving on the committee is probably the best person to offer that training.

Each member needs a healthy skepticism about claims; a drive to search for supportive information; a

concern for impartiality; a sense of the necessity for objective information as a mold for and a brake on subjective opinion; an insight into the differences in information sources which can affect the validity, the reliability, and the comparability of information; and a realization that goals govern evaluation design, together with an appreciation that several different goals may all be worthy goals. Such a realization should underlie every task the committee undertakes, from the establishment of criteria to the final selection.

Teachers and Selection Committees

Service on selection committees is too often something teachers are expected to do in their spare time and perhaps at some expense to themselves. Selection is too important to become the victim of such attitudes, and until it becomes an equal partner with other demands on the teacher's time — teaching, preparing lessons, and conferring with parents and students — it must be an activity for which the school makes arrangements, seeing that the selector's classroom is covered by a substitute and that his expenses are paid.

Sharing the selection task can be extended even further. The selection process will be much less burdensome if groups of schools share the time-consuming job of reviewing materials in the light of generally applicable criteria. The final selection, of course, must be made by each individual school.

Committee leaders will find the selection process smoother and more efficient if they will spend some time with the committee discussing how selection should be accomplished, and if a practice evaluation session is scheduled. Such sessions make it possible to resolve differences in how to approach the task before work begins.

Establishing Criteria

After the selection committee is organized, oriented, trained, scheduled, and informed about the goals of selection, criteria by which materials will be selected should be specified. A systematic statement of criteria, cooperatively developed, is a useful tool. The exercise of developing the statement actually helps to make the bases for selection explicit to all.

Use of the statement as a basis for developing evaluation procedures and tools can ensure thorough evaluation of each item under consideration, and it permits selectors to defend their choices at review time. Moreover, successive statements supply subsequent selectors with a valuable historical record, particularly useful when change is contemplated.

Criterion statements are likely to be long and detailed. They must cover curricular requirements, pedagogical requirements, administrative requirements, and demonstrated effectiveness requirements. They are, for the most part, detailed explications — again, in concrete behavioral terms — of what materials must make possible for students and teachers in order to meet the specified goals.

It is becoming more and more necessary that goals be hammered out, written down, and made explicit to the school's staff and its public. Certainly they are basic to all school planning, not just materials selection; and, in this day of increased interest and involvement in public affairs by all manner of groups and individuals, it is valuable to be able to publicize just where one stands on certain basic points before anyone feels it necessary to ask.

Levels of Criteria

Several levels of criteria can be identified. The simplest level, which might be called *descriptive* criteria, concerns such matters as course and grade level, cost and other purchase constraints, recency of the material, availability of assistance in implementing the material, and coverage of subject matter (for example, whether square root is included — *how* it is covered is a point which must be considered in detail further along in the selection process).

Another level of criteria has to do with the *instructional setting* in which the material will be used. Questions of pupil, teacher, school, and community characteristics (all of which must be accommodated comfortably) must be answered: Is the material designed for pupils whose interests and achievements match those of the students for whom materials are being selected? What about their socioeconomic, geographic, and ethnic orientations? Are the demands on the teacher too great for one who has a number of other types of lessons to prepare for, but interesting and challenging for one who teaches only this material to several groups of students? Does the material require physical facilities and scheduling which are not possible in this school?

Then there are more subtle criteria — questions of *aims, approaches, strategies, methodology, rationale, point of view, scope,* and *pacing.* A valuable exercise early in the selection process is a thorough discussion of potential criteria to make sure that the members of the selection committee agree about what they are looking for — these are not cut-and-dried matters. Will these arithmetic materials merely demonstrate the technique of deriving a square root, or will they encourage students to discover how to do the job? Do these general science materials put too much of a premium on

reading? Are these materials so organized that a relatively inexperienced teacher who is not a specialist in the field can help students use them effectively? A text that encourages student exploration will be of little use to teachers who wish, or are expected, to direct student activity very carefully. In contrast, students looking for ideas for the Westinghouse Science Talent Search would not find much in a film designed to explain photosynthesis to secretarial students.

For clarity and usefulness, all *outcome* objectives (and just as many descriptive facts as will fit) should be stated in behavioral terms that can be translated into points to look for when examining materials — what a person can do, not what he thinks or feels. For example, if an outcome objective is that a student should be able to describe the steps Congress takes in resolving indecisive Presidential elections, this objective can be translated into points to be looked for when examining materials. On the other hand, if the objective merely states that a student must comprehend the significance of the relationship of the Congress to the Executive Branch, it is not translatable to concrete points to look for.

Even less measurable are the all too frequent statements of the following sort: "The aim of the course is to establish the student as a good citizen by inculcating in him the traditional values of his forebears." Probably no way exists for a textbook or film to accomplish that job, and certainly it is not translatable to specific points to look for in evaluating instructional materials.

What characterizes behavioral statements of goals is that they say what a person must be able to do about, with, or as a result of the materials under consideration: "The student can describe . . .," not "the student understands. . . ." "The student can contrast . . .," not "the student appreciates. . . ." The very act of devising

and specifying such goals can be valuable in defining and clarifying just what it is one wants to do.

Effectiveness is the final, most important, criterion — whether the materials demonstrate empirically that they fulfill the goals set for them. *Opinion* about effectiveness is helpful, but it is rarely conclusive. What is needed is *actual evidence* that students' behavior changed in the desired direction as a result of having used particular materials.

The selection committee will want to call on an evaluation specialist to formulate the standards of effectiveness materials must meet in order to be retained in the pool of materials to be considered; this is one of the technical matters which makes his presence on the committee essential.

Criterion Check List

This check list, designed to be suggestive rather than exhaustive, represents how a school system might organize its criteria. The list is not intended to dictate any particular organization.

I. ADMINISTRATIVE REQUIREMENTS
- A. Schedule
  1. fit to school-year organization, including possibility for adaptation
  2. fit to school-day organization, including possibility for adaptation
  3. lead-in time needed for pilot use, training, etc.
- B. Budget
  1. initial cost for materials and training
  2. ongoing costs for expendables, replacements, additional training, etc.
- C. Personnel
  1. additional specialized personnel

2. additional nonspecialized personnel
3. schedule and budget for training as required

D. Space

E. Legal Constraints

II. CURRICULAR REQUIREMENTS

A. Age, Grade, Ability Level

B. Subject Matter Content

1. selection and arrangement
2. scope
3. sequence of presentation
4. point of view, including treatment of minorities, ideologies, sex roles, etc.
5. media of presentation
6. necessity for supplementary materials to complete presentation

III. PEDAGOGICAL REQUIREMENTS

A. Instructional Setting

1. suitability of socioeconomic, geographic, ethnic orientation
2. student interest, achievement, learning style
3. teacher characteristics
4. suitability of physical space for recommended implementation
5. fit of recommended presentation into standard schedule

B. Teachers

1. necessary background
2. pedagogical style
3. teaching schedule and load and other school responsibilities
4. provision for special training
5. necessity of supplementary personnel
   a. professional
   b. lay

C. Methodology
   1. fixed or flexible?
   2. specified in detail or implied in materials?
   3. compatible with a view of how learning takes place?
   4. pupil-centric or teacher-centric?
   5. individual, small-group, large-group, or some combination?
   6. if individual, programed?
   7. heuristic or didactic?

IV. EVALUATION REQUIREMENTS

A. Method for assessing outcomes of instruction and learning
   1. self- or teacher-administered and scored?
   2. descriptive or objective?
   3. norm or criterion as reference?
   4. diagnostic?
   5. prescriptive?
   6. frequency of measurement and reporting

B. Producer's empirical evidence
   1. developmental testing reports
   2. reports on field testings of final form of materials
   3. evidence of collection of feedback on materials in use

C. Reports from other schools
   1. statistical
   2. anecdotal
   3. direct, or through [illegible]ganization like EPIE*

D. Results of local pilot use
   1. statistical
   2. systematically observed
   3. collected by questionnaire

* Educational Products Information Exchange Institute, 463 West St., New York, N.Y. 10014.

a. from teachers
b. from students

4. anecdotal
   a. from teachers
   b. from parents
   c. from students
   d. from other sources

E. Results of follow-up during operational use
   1. statistical
   2. systematically observed
   3. collected by questionnaire
      a. from teachers
      b. from students
   4. anecdotal
      a. from teachers
      b. from parents
      c. from students
      d. from other sources

Examination and Review

Examination and review on the basis of a criteria check list must not be considered the final step in the process for selecting instructional materials. Comparing materials with criteria is important, but the comparison gives little or no *direct* evidence of whether materials are effective. *The purpose of examination and review should be to select materials for further, empirical, evaluation.*

We have assumed in the past that materials examined and judged acceptable as to content and approach will also be effective learning tools. That may be true, yet research indicates that examination of materials is an unreliable way to judge their learning effectiveness.

For instance, in one study a principal and a group of teachers were asked to rate for learning effectiveness alternate versions of materials for which empirical evidence of effectiveness had been collected but not disclosed. Their ratings had a correlation of —.75 (yes, that's a minus) with the empirical measures.

A replication of the study by another researcher produced similar results. But we have not taken much notice of findings of that kind; we are often impervious to research results, particularly when they relate to an area in which we have not yet recognized the existence of a problem. Once we recognize the problem, however, we must see that we use examination and review as one step, not a final move.

Meeting the demands which the greatly-expanded supply of materials puts upon us, our examination-and-review procedures should be refined and systematized.

To ensure that examination is thorough and that observations are reported in comparable terms, the selection committee may want to devise evaluation check lists, questionnaires, or, if professional evaluators are available to help, carefully constructed and weighted rating scales. Good forms greatly facilitate the selection process; they make it possible to summarize and compare committee opinions.

Check lists, questionnaires, and rating scales ensure complete coverage of relevant points about materials. They also help to systematize the collection of opinion, and provide useful records for reference. The selection committee will benefit by translating its statements of criteria into a set of appraisal forms, a series of hurdles which materials must negotiate in order to stay in the running for final selection.

The committee will find, when it comes to writing all these appraisal forms, that it will not be too difficult if goals and criteria were worked out carefully. Schools, libraries, curriculum centers prepare them all the time; here are some hints, based on a review of ones which have already proven themselves:

- The most useful forms are designed to help the user make a good appraisal, not leave him floundering about, wondering what he should look for and what he should do with what he sees. For instance, instead of merely asking for a check mark to indicate whether a teacher's manual specifies a methodology, a useful form will ask about specific aspects of methodologies: Are lessons planned in time sequences? Are suggested procedures easy to follow? Is there special emphasis on probable problem areas? Are supplementary drills, discussion questions, and outside activities suggested?

- One might stretch the appraiser's perspective by asking him to respond in the context of comparison: Is it important to buy the color version of this film? Will this tape hold the student's interest better than a lecturer would? Is this new film so far superior to the one on hand that it should be substituted for it?
- The committee should provide for some indication of the relative importance of the criteria. After careful thought and discussion, priorities can be assigned and some flexibility allowed for dealing with priorities so that trade-offs can be made. A surprising and totally unexpected feature may turn up to justify ignoring some previously important criterion. The committee should be reminded about the value of stating criteria in behavioral terms so that there will be points to look for in checking out instructional materials, to see whether they are really designed to help students learn what they are expected to learn. If the criteria are stated behaviorally, that part of the appraisal form writes itself, so to speak.
- In the quest for a rational system and organization, one sometimes forgets the vital importance of open-ended comment — for successful appraisal *and* for the responder's morale. For instance, if a teacher is asked to rate a film from "superior" to "poor" on its capacity to stimulate classroom discussion, how will he rate the film if it presents its case *so badly* that discussion, paradoxically, will be full, lively, and productive? A little "Why?" or "How?" under the rating scale would help him show just how the film would fit into the instructional scheme.
- A special word of caution: appraisal forms are not tests, and scores should not be the end results. If

scores were used, two books might receive the same score, one possibly because it is attractive visually though quite inaccurate as to facts; the other possibly because it is accurate but also badly printed, ugly, and difficult to read.

- The selection committee should prepare instructions for its members to follow. Almost any point is open to more than one interpretation, and it is important that the committee come to agreement on each point and take steps to ensure that members bring the same understandings to their examinations of materials and their integration of information about those materials. Only then can results for many instructional items from several committee members be compared.

Basic Descriptive Information

The obvious source of basic descriptive information about instructional materials is the producer. Although many of the big publishers, producers, and manufacturers alert schools about their educational materials, selection committees cannot assume that they have the necessary information on all the materials that exist.

Making certain that the materials to be evaluated are a complete collection is important. Such a list will prevent a school's missing good items produced by small companies with small sales forces and small advertising budgets.

Several sources of descriptive information are listed here.

- Directories or collections of catalogs provide a good place to start. Some of them, it should be noted, are supported by producers and are less than complete and unbiased.
- Reviews and listings appear in general periodicals, in professional journals, in yearbooks, and in special bibliographies.

- "Guides to the guides" supply listings, sometimes with descriptions, of catalogs, bibliographies, and periodicals that are useful in the selection of library books and nonprint media materials.
- Educational Products Information Exchange (EPIE) supplies descriptive information also, in a format much easier to use than separate pieces from each producer. EPIE listings contain tables of comparable information about all the items, from all producers, available for a given purpose. One bit of information included is whether the developer will supply field-test or other study data about his product.
- An increasingly common source of information, particularly for supplementary materials, is a central collection of materials — an instructional materials center, an educational media selection center, or similarly named facility — where one may examine materials and discuss them with a curriculum specialist.

Descriptive criteria can be expressed in concrete terms and therefore used to make the first quick cut through the list of available materials so as to reduce it to a manageable size.

Effectiveness Criteria

Effectiveness information from sources outside the school — the producer, or other schools which have used the materials — is rarely available. When it is, it can be used early in the selection process. Information on *how many* and *what types* of students the materials were used with, the *method* of measuring the materials' effectiveness, and the *results* of the measurement can be very revealing, when reliable and if used with care.

Evidence that a specific instructional material really does its job can outweigh any factors giving some

members of a selection committee reason to doubt its effectiveness. It is important that the evidence be pertinent to the goals and the conditions of use guiding the committee's choice. Rarely will the conditions under which a producer tested his materials, or a school used them, match exactly those in the school making a selection. The differences in students, teachers, and conditions must be noted and assessed; they may modify one's reliance on the resulting information.

The more reports there are, of course, the more chance there is for a match, and similar reports from disparate sources might justifiably be extrapolated to one's own situation.

Selection committees would do well to seek this evidence of effectiveness vigorously, and to give priority to materials which are supported by it. As more schools recognize and require this kind of evidence, more and more sources will come into view. Sometimes producers can supply their own data and also names of schools who have used their materials and are willing to share their information. Sometimes one school knows through channels that certain other schools have used materials under consideration. As the body of pilot-use and effectiveness information grows, the exchange of such data among schools, carefully described as to population, technique, and instruments can be facilitated.

Selection committees will save time, money, and effort if producers would point out with what groups their materials have been successful. EPIE's research, as reported in testimony to the U.S. Congress in May 1972, indicates that only about one percent of the 200,000 materials on the market have been subjected to any systematic, scientific investigation, by the producers — whether they do their job and with whom they do it best, if they do it all. The reason, said one

producer, is that school people have not demanded such evidence.

Content and Coverage Criteria

When it comes to applying the subtle criteria of content and coverage, methodology and rationale, a simple check list is not likely to suffice. The committee will want to devise a more probing form, perhaps a rating scale or a questionnaire which permits recording of observations. Members of the committee will want to study the materials in detail to determine how they measure up to these subtle criteria, even if there is a detailed analysis of a potential material available from an impartial source.

Sometimes the testing of the materials against these criteria is done individually by committee members, sometimes by the whole group in consultation with producers' representatives. Whichever way, the evaluations should lead to discussion. Inevitably, materials still under consideration at this stage will require some joint committee attention to ensure agreement about how they fit into the school's program and what steps remain to reach a final decision.

Pilot Use and Testing

Eventually, the selection committee gets the original list reduced to a few promising materials. That is when pilot use and testing should be initiated. It is important that the selection committee decide just what questions they want the trial use and testing to answer, and just how good the materials must be to be selected — in other words, the effectiveness criteria which materials must meet. The committee's questions will govern the nature of the trial procedures.

No more effective method of judging materials has ever been devised than actual use with students. Even under simulated conditions, as in a selection workshop, it is possible to see differences between the results of very careful analysis and those of actual trial.

In one such workshop, participants assessed curriculum materials and examined a complex educational game dealing with ecology. They worked in groups simulating selection committees and developed detailed analyses of the game, based on instruction booklet information and observation of the playing board and pieces. They then began to play the game — and to change their ratings of it and of its explanatory material.

Instituting the pilot use of materials in a school, with proper provision for pre- and post-testing and for matching control and experimental student groups and teachers, is a task for specialists; it can only be described briefly here. (Two pilot studies which have unusual aspects and have had high impact are reported in Appendix A.)

Pilot use and testing give the school system empirical evidence of how well an educational item will teach. A plan can be elaborate enough for a whole cur-

riculum or simple enough for a single-concept film loop.

In typical cases, a plan can use trial materials with one group of students and well-established ones with a comparable group, and then compare the performances of the two groups on the same measure of "criterion mastery" — perhaps a commercially available standardized test of the skill or subject being taught, or perhaps a locally developed test that is kept independent of the two learning sources.

In other cases, tests of knowledge or ability are given before the students study the new material *and* after, to see what they have learned — if anything.

An achievement test for media which measures the extent to which various attributes of films, filmstrips, and other such materials influence desired or designated behavior is also available from EPIE. Attributes like length, target grade level, vocabulary level, skill level, achievement gain, and command of attention are comparably measured for various items and plotted into profiles that greatly facilitate comparison among those items. The most elaborately devised index is that of achievement gain, which is based on pre- and post-test use of questions designed to measure the extent to which the instructional item meets a given behavioral objective.

Pilot students and teachers must be carefully described, and outcome behaviors must be carefully defined. If they are, pilot use and testing can show, for instance:

- that a given book is very effective in a small, highly motivated class taught by a Ph.D. in the subject matter but that it is much too hard for the typical student to understand
- that the beautiful book, full of four-color illustrations, maps, and cut-outs to use in model-making, does not impart basic information as well as the

small brown textbook with the closely marshalled facts and the straightforward presentation

- that a certain film adds to a slow student's knowledge because he is allowed to view it on his own schedule and stop and repeat portions as he wishes
- those areas in which the materials do a particularly good, or a particularly bad, job
- how to help in the appraisal of a particular methodological approach.

In pilot testing, *how* teachers and students use the materials should be the subject of systematic analysis. Equally important is *what* users think about materials. Questionnaires for this purpose should give pilot users opportunities for open-ended comment. Students should have the chance to respond anonymously.

Implementation and Follow Up

When all the evidence is in, the committee will be ready for a decision. They may be fortunate to have clear-cut evidence of what to do, but it is more likely that weighing and sifting will be needed. Often the committee will be required to account for its selection to the public or school authorities or both. Organizing information for decision making will also organize it for such eventualities. The effect which public or authorities can have on final selection varies from situation to situation. Usually, carefully laid groundwork and thorough procedures make acceptance automatic.

Another outcome of the process can be the sharing of information with others who would find it useful. A good idea, too, is to offer the information to the producers of the materials tested. They may also be interested in summaries of responses to appraisal forms, to show what the committee found or failed to find in the materials to satisfy the demands they were making on them.

Empirical information of this kind should have a significant effect on the producer's plans for revision and

even perhaps on his overall approach to the development of instructional materials. Of course producers need information from a variety of schools, but carefully developed information for one school with clearly defined goals and conditions will be an enormous help.

In order to implement its choice under favorable conditions, the committee will want to give parents, students, and teachers an opportunity to hear about it and ask questions about it. Pilot-use teachers, students, and parents can be helpful in this orientation task. The most extensive and formal approach must be to the teacher, who may have to spend considerable time and energy in training sessions, workshops, and lesson preparation. Frequently the producer includes training as part of his instructional package.

Once selection has been effected and materials are in regular use, it will be valuable for future selection to check out how successful the selection was. Similar sorts of controlled testing as that described for trial use will reveal on-the-job success or failure — or, rather, degrees of each — and assist in improving selection procedures for the future.

The committee will want to know how their selections work out and whether students do indeed learn from the materials selected. If not, of course, materials must be changed or their use modified. But maybe selection procedures can be improved, too: What points did the committee miss which would have indicated the inadequacies which showed up in operation? What parts of their procedures led to their error?

Follow-up testing, observation, and the gathering of student and teacher opinion are tried-and-true methods here as well as in pilot use and testing. The ultimate goal is good learning; vital to success is good material; vital to good material is revision and refining and planning based on how well material meets the goal.

What Can We Expect Others To Do?

Changing our well-established, comfortable behavior will be difficult. What will be more difficult will be to persuade producers to change their well-established, comfortable behavior.

The first step to doing this is to get agreement that these behaviors can and must be improved. What must be avoided at all costs is, on the one hand, preaching a counsel of perfection — "Research use cannot match operational use exactly to the last detail, so let's not attempt it" — and, on the other hand, becoming defensive about established practices: "These practices have been developed and refined through professional experience over the years. We need no ivory tower revision of our procedures."

What we must try to do is get all education professionals, in industry and in schools, to subscribe to the necessity for improvement and accomplishment — to say: "There isn't a product that cannot be improved. Every product, new and old, must continually be revised in light of growing knowledge and the constantly changing needs of learners."

Such continual evaluation and revision works for the producers as well as for teachers and students. Quality materials sell because they do the job well. Let's look at the introductory college text in economics that for many years has outsold all others in its field.

More than 20 years ago Professor Paul Samuelson of the Massachusetts Institute of Technology wrote a text that now appears in its eighth edition. Each edition has gone through a three- or four-year revision cycle to update content, organization, and style of presentation. To

prepare for each revision, Professor Samuelson gathers information from both teachers and students regarding the teaching-learning effectiveness and the acceptability of the text.

Similar approaches must be taken in the development of materials before they are made generally available as well as in revising them for subsequent editions. It is beginning to happen.

A decade ago the federally-funded Biological Sciences Curriculum Study Group developed three sets of biology materials, for three different approaches to biology learning. All the materials were tried out thoroughly on appropriate groups of students and revised accordingly before they were made commercially available.

The Southwest Regional Laboratory for Educational Research and Development, another federally-supported group, has completed a *First Year Communications Skills Program*, 10 units designed for kindergartners, for which they have done extensive developmental testing. The producers say: "The reading skills to be acquired are listed very explicitly throughout the program and their successful mastery is the most important result of the instruction." That language is similar in tone and content to the language one finds in connection with commercially developed materials, but the kind of evidence offered by the Southwest Laboratory in support of its statement is hardly ever forthcoming from a commercial developer.

Back-up data were gathered over a four-year cycle of product development, during which the units were tried over and over with kindergarten children and continually revised. The developers state categorically that criterion mastery was not achieved when the first version of the product was used with children. But now, after several years of continual evaluation and revision, the effectiveness of the product for its criteria has been

verified. It has been tested with 30,000 children in 12 states, in middle-sized schools, and in several large urban districts, and the materials are ready for commercial distribution.

One type of instructional product subjected for many years to pre-publication testing and revision is the standardized test. Producers of standardized tests administer experimental forms of the tests to samples of students in order to determine test reliability, difficulty, time and speed, and such. They then revise and rearrange questions in light of their findings in order to produce a final test.

The final test is administered to a carefully described student population in order to determine its validity for measuring the performance of such students (and perhaps to establish norms against which other students' performance can be compared — an extra and quite complex wrinkle required by the nature of the use of tests). In many cases, data are collected from users of the final test to assist in refining the next edition.

All educational products should be developed and revised in this manner, on the basis of actual student performance, though of course conditions of use for other materials cannot — and need not — be as rigorously controlled as must be the case with tests.

Depending on the nature of what the material is designed to teach, the population of students who try it and then take tests on what they have learned may be divided many ways: by age, ability, sex, geographic location, socioeconomic background, professed interest, etc.

How the materials are used — the teacher's background, training, and methodology; class and other time devoted to the materials; other materials used in conjunction with them — can in some cases be prescribed and in all cases described.

The results of such testing could be not only excellently articulated instructional materials but also usefully detailed descriptions of just how and with whom to use the materials. If schools begin to demand evidence of effectiveness as well as expert opinion and testimonials from satisfied users, producers may find that they are producing fewer, but better-selling materials.

> "A couple of years ago, we wanted to do some field testing, but scheduling wouldn't allow it. . . . It takes too much time and we wouldn't have gotten the books out. . . . It wasn't a question of money, but just scheduling. We're now working on a program we plan to field-test. I hope we can. . . . Testing has lots of problems, you know. . . ."
>
> *Vice President and Editor-in-Chief of a major publishing company*

How do we get the producers to begin?

They know that learner-verification will take time, money, and cooperation from schools. They will be reluctant to add the necessary time to their schedules, but we can be certain that they will do so if only in that way can they develop materials acceptable to those who buy their products. They will rightfully expect cooperation from schools in supplying and describing trial populations and in meeting their specifications for the use of the materials. If increased costs are too much for the producer and the purchaser (to whom they normally would be passed) to absorb, federal funds might be forthcoming, as research and development grants for producers and as financial aid to schools which agree to buy only learner-verified materials. (See Appendix B.)

With his problems of time, money, and cooperation well in hand, how does the producer proceed?

He may want to add evaluation specialists to his staff, to plan research, and to attend to such matters as selecting samples and instruments, treating and interpreting data, and describing the population and the

teaching conditions. He will need, too, to find out just what it is purchasers want to know — guidelines for initiating his developmental testing program — and minimum standards his materials will be expected to meet. These guidelines and standards would be best developed, perhaps under the federal aegis, perhaps by the National Institute of Education. Here, again, cooperation from schools and from producers would be needed, as well as from evaluation experts.

The producer may want to report his findings in technical manuals such as those described for tests. A technical manual might provide answers to specific questions such as these:

- Can students who could not repair carburetors before they saw the film on the subject do so afterwards?
- Did testing show that the use of a female mechanic in the film actually facilitated learning?
- What kinds of students learned to repair carburetors — those with no other training in auto mechanics as well as the relatively experienced?
- Can the film be used independently, or must an experienced teacher guide students through it?
- How much is learning increased if students have access to an actual carburetor while viewing the film?
- Did the groups who saw the film learn more than similar groups who heard lectures?
- What about lecture and demonstration — does the film do better than that with experienced students?

. . . And so on, as we seek for materials that match our criteria.

In addition to empirical evidence of performance with carefully described populations under carefully described conditions, the manual should describe the objectives that the material is designed to meet, the

rationale underlying its design, the types of students and teachers toward which it is aimed, and the qualifications of the educators who developed it. The manual can also be used to inform prospective users of what success other users have had.

Until now producers have been curiously uncurious about results of studies involving their materials, even when the results were favorable. A school official in a large Midwestern city that engages in extensive pilot testing of material under carefully controlled conditions before buying in quantity has said, "We couldn't force our data on . . . [the producer]. He just wasn't interested." Yet, other school systems, and not just those in large Midwestern cities, could learn from those results a great deal about using the materials in question.

Conclusion

To enhance learning by means of using the best possible instructional materials, we must make demands upon ourselves and upon our colleagues in the education industry. We must, above all, require a systematic, scientific approach to the appraisal of materials, to see whether and with whom they work. We need products of proven worth in which producers and school people alike can have confidence. The ultimate consumer, the student, has a right to expect no less from us.

Appendix A

I. The Media Center as Curriculum Resource

A medium-sized school system on the East Coast has just about abandoned the use of single textbooks as a somewhat surprising result of its plan to make libraries available in every elementary school and to give in-service training in library-media center use. Implementation of the plan prompted teachers to ask to review the school system's curriculum. Accordingly, a plan was established to do just that and to install new curricula, under the supervision of the school system's Educational Development Center.

Under the plan, research committees now investigate current curriculum trends to define basic objectives and to work out the scope and sequence of topics to be covered. Then a larger faculty committee defines a program in more detail and makes suggestions for implementation, grade by grade. In due course a "pilot school" tries the program for a year. A year-long program of workshops ensures that all teachers are informed about and contribute to the new curriculum development and keep up with suggestions as to methods, media, services, and activities. After the pilot year, the new course is introduced to the whole school system with whatever revisions it requires — or abandoned, depending on the outcomes.

One basic outcome of this cooperative curriculum development has been the substitution of a variety of books and other media for the single textbook. Media for use in implementing the curricula are basically teacher-selected. Suggestions are made by committees and by pilot school faculties, and sessions are held to

introduce teachers to new media available. But, in the long run, the teacher is the final arbiter of what will be used in his classroom.

The school system maintains a large collection of media for teachers to choose from. Most of the items are housed in individual school libraries but centrally cataloged to allow system-wide use. A teacher may ask to have the collection augmented to fit his needs. Indeed teachers are encouraged to suggest additions to the media collection. A good start is made during the pilot year. After new curricula are in operation, teachers may ask for materials for review, and they may collect student opinion of the item as well as judge it for themselves. They record preview results on a form which the Educational Development Center files for use by teachers who later may consider reviewing the same material. In the case of films, it is sometimes possible to have system-wide "film shows" where concerned teachers can see and make comments on numbers of films. The basic library collections are maintained by library specialists, but special books are added by a process similar to that for media.

II. Shall We Try This Medium?

Not all pilot use has to do with content. It is also a fine method of discovering whether a medium is an effective teaching tool for the group one is responsible for. An instructional resources center in the Southwest did a careful trial run to see whether to add Super 8mm film loops to its collection. They asked the users — teachers and students in five schools — to comment on the effectiveness of the medium in general, not the content. A large collection of loops was lent by producers for a whole semester, to ensure a thorough trial.

The trial-run teachers had had no experience with film loops and had little idea of what to expect from

them. To give them some background, the study director supplied descriptions to help them categorize the loops into the "inquiry" class, the "demonstration" class, the "physical skill developers" class, and the "cognitive skill developers" class. He suggested applications for each type — small group seminars for one, individual practice for another, concept clarification for a whole class for a third. He also told the teachers exactly what questions they and their students would be asked about film loops as a teaching and learning device when the semester was over.

The questions were written in recognition of the fact that the user's opinions would be subjective, not "scientific." Student ratings dealt not only with the overall impression but also with whether the loops made them think or bored them, whether the absence of sound bothered them, whether the loops should be used in classes or put in the library to circulate. Degrees within the rating scales for students kept their vocabulary in mind:

5	4	3	2	1
I liked them very much	I liked them at first but as the novelty wore off not so much	They were OK	I didn't understand them	They were no good

Questions to teachers were considerably more comprehensive. They had to report on how and how often the loops were used as well as rating their "personal reactions" to the effect of the loops on their classes. They were asked to rate how successfully the loops made their points and whether the system operated smoothly. Finally they were asked to answer in writing these questions about the possible impact of loops on the status quo:

1. Did the loops require a change in your teaching techniques? If so, how?

2. In your opinion, did the results justify the change?
3. Do you feel that the film loops offer more teaching possibilities than filmstrips?
4. Is the lack of sound valuable in allowing you and your students to discuss freely what is shown, or is lack of sound a drawback? Please comment.
5. If the film loops are comparable in price to filmstrips, would you feel that the loop system would be a valuable addition to your teaching?
6. Do you feel that groups of loops, arranged in kit form, should be a circulating item in the regional library, or from your experience would they lose value if they had to be ordered 10 days in advance and kept only one week?

On the basis of responses to the survey, film loops have become a much-used part of the central instructional resources collection.

Appendix B

The Detroit Study of Treatment of Minorities in Textbooks

A special case of systematic evaluation and review is worth reporting, not only because it comes to grips with a problem of increasing importance but also because it is an example of successful cooperation between schools and producers.

In 1964 the Board of Education of the City of Detroit instituted an additional step in its textbook selection procedures. After a selection committee makes its recommendations, the Department of Intergroup Relations makes careful analyses of recommended books in terms of 20 criteria developed by the department. The criteria require positive (rather than mere absence of negative) treatment of minority groups. The evaluators are required to say whether, in its text and its illustrations, a book or curriculum material shows fully integrated groups, shows members of minority groups in positions of leadership, analyzes intergroup tensions and their causes fairly, and shows the contribution of members of minority groups of American life, to name a few. The results are reported to the superintendent of schools for his use in recommending books for adoption to the board of education.

Since so few of the textbooks available were found to meet the criteria used in evaluating books for the Detroit Public Schools, the staff invited representatives of major textbook publishing companies to meet with them to discuss the situation. According to the report on the procedure prepared for the U.S. Senate Permanent Subcommittee on Investigations, the meetings "appear to have been promising and fruitful. Promising, because

some publishers indicated that their meeting was most helpful '. . . in shaping our thinking for the days ahead.' Fruitful, because a number of publishers' representatives offered to make specific changes and improvements in their textbooks. Some of these have already been made."

Bibliography

1. Center for Instructional Research and Curriculum Evaluation. *The Accountability Notebook*. Urbana: Center for Instructional Research and Curriculum Evaluation, University of Illinois, 1971.
2. Chauncey, Henry and John E. Dobbin. *Testing: Its Place in Education Today*. New York: Harper and Row, 1963.
3. Educational Products Information Exchange Institute. *Evaluation Practices Used in the Selection of Educational Materials and Equipment*. Albany: Division of Educational Communications, The State Education Department, The University of the State of New York, 1969.
4. Magers, Robert F. *Preparing Instructional Objectives*. Palo Alto: Fearon Publishers, 1962.
5. Miller, Richard I. *Selecting New Aids to Teaching*. Washington: Association for Supervision and Curriculum Development, NEA, 1971.
6. *Selecting Instructional Materials for Purchase*. Washington: Joint Committee of the National Education Association and Association of American Publishers, 1972.
7. Tyler, Louise L., M. Frances Klein, and William B. Michael. *Recommendations* for Curriculum and Instructional Materials. Los Angeles: Tyler Press, 1971.